

**Remarks**

Reconsideration of this Application is respectfully requested.

Upon entry of the foregoing amendment, claims 1-5, 8-13, 16-19, 22-26, and 29-50 are pending in the application, with Claims 1, 16, 29, 39, 49 and 50 being the independent claims. Claims 6-7, 14-15, 20-21 and 27-28 were previously cancelled. No new claims are sought to be added.

Based on the above amendments and the following remarks, Applicants respectfully request that the Examiner reconsider all outstanding rejections and that they be withdrawn.

***Arguments in Previous Office Action Response and Examiner Interview***

Applicants once again extend their appreciation to the Examiner for meeting with Applicants' counsel on November 10, 2004. During that meeting, Applicants' counsel explained the invention with a particular focus on independent claims 1 and 29. In particular, Applicants' counsel explained how a segment-level actual usage value for one or more word combinations can be determined, and also explained how a segment-level expected usage value for one or more word combinations can be computed.

During the Interview, Applicants' counsel demonstrated that U.S. Patent No. 6,678,694, entitled *Indexed, Extensible, Interactive Document Retrieval System*, issued to Zimmerman, et al. on Jan 13, 2004 ("Zimmerman Patent") and U.S. Patent No. 5,745,776, entitled *Enhanced Electronic Dictionary*, issued to Charles Sheppard, II on Apr. 28, 1998 ("Sheppard Patent") did not render the claims of the present application obvious under 35 U.S.C. § 103(a).

Applicants provided an Amendment and Remarks, dated November 12, 2004 ("Nov. 12<sup>th</sup> Amendment and Remarks"), in response to the Office Action received from the Examiner, dated July 14, 2004 ("First Office Action"). In the Nov. 12<sup>th</sup> Amendment and Remarks, Applicants further demonstrated that the claims, as amended, should be allowable over the references cited by the Examiner in the First Office Action. Specifically, the Examiner had provided rejections under 35 U.S.C. § 103 based on various combinations of the Zimmerman Patent, the Sheppard Patent, U.S. Patent No. 6,446,061, entitled *Taxonomy Generation for Document Collections*, issued to Doerre et al., on Sep. 3, 2002 ("Doerre Patent") and U.S. Patent No. 6,038,560, entitled *Concept Knowledge Base Search and Retrieval System*, issued to Kelly Wical on Mar. 14, 2000 ("Wical Patent").

Applicants provided a second Amendment and Remarks, dated August 9, 2005 ("Aug. 9th Amendment and Remarks"), in response to a second Office Action received from the Examiner, dated February 26, 2004 ("Second Office Action"). In the Aug. 9th Amendment and Remarks, Applicants once again demonstrated that the claims, as amended, should be allowable over the references cited by the Examiner in the Second Office Action. Specifically, the Examiner had provided rejections under 35 U.S.C. § 103 primarily based on combination of the Zimmerman Patent and U.S. Patent Application Publication No. 2004/0024583, entitled *Natural-Language Processing Using a Large Corpus*, filed by Robert J. Freeman on March 20, 2001 ("Freeman Application").

In the Present Office Action, the Examiner has introduced a new set of three references (identified below) along with the Zimmerman Patent as the basis for additional 35 U.S.C. § 103 rejections. While the Examiner did not specifically address the Applicants' previous arguments supporting allowance, in light of the Examiner's reliance on a new set of references and new bases for additional non-final rejections, Applicants conclude that their Aug. 9<sup>th</sup> Amendment and Remarks were persuasive and, like past amendments and remarks, have overcome all previously outstanding rejections. Applicants now turn their attention to the new rejections in the present Office Action.

***Rejection Under 35 U.S.C. § 101***

Claims 1 and 29 have been rejected under 35 U.S.C. § 101 because "a method to identifying topics in a data corpus having a plurality of segments" and a method to display a list of topics associated with data items stored in a database are "descriptive material." Present Office Action at 2. The Examiner contends that "the method is not being tangible to form the basis of statutory subject matter under 35 U.S.C. § 101."

Applicants respectfully traverse this rejection. Both of these claims produce useful, concrete and tangible results that have "real world results." In the case of claim 1, a topic is produced that serves as a useful, concrete and tangible result having "real world results." For example, a topic can be associated with a document to later facilitate efficient retrieval of the document. In the case of claim 29, a list of topics is displayed that serves as a useful, concrete and tangible result having "real world results." For example, the list of topics can be used to narrow a search field and more readily identify a document to be retrieved.

The USPTO's Interim Guidelines for Examination of Patent Applications for Patent Subject Matter Eligibility ("Interim Guidelines") specify that tangible is defined as something that is or produces a real world result, but is not limited to a machine or apparatus. Interim Guidelines at 21. As discussed above both claims 1 and 29 produce real world results that have practical applications and uses. Thus, claims 1 and 29 clearly recite statutory subject matter. Applicants respectfully request that this rejection be reconsidered and withdrawn.

***Rejections Under 35 U.S.C. § 103***

Claims 1-11, 13, 16-24, 26 and 49 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 6,363,378, entitled *Ranking of Query Feedback Terms in an Information Retrieval System*, issued to Conklin *et al.* on Mar. 26, 2002 ("Conklin Patent") in view of United States Published Patent Application No. 2004/0128267, entitled *Method and System for Data Classification in the Presence of a Temporal Non-Stationarity*, filed by Berger *et al.* on May 10, 2001 ("Berger Application"), and further in view of United States Published Patent Application No.

2002/0103799, entitled *Method for Document Comparison and Selection*, filed by Bradford *et al.* on Dec. 5, 2001 ("Bradford Application"). Applicants respectfully traverse these rejections and request reconsideration.

***There is a Lack of Motivation to Combine the Conklin Patent, the Berger Application and the Bradford Application.***

At the very least, there is no teaching, suggestion or motivation to combine the Conklin Patent, the Berger Application and the Bradford Application to render claims 1-11, 13, 16-24, 26 and 49 unpatentable. Obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either explicitly or implicitly in the references themselves or in the knowledge generally available to one of ordinary skill in the art. MPEP 2143.01. The Examiner has provided no permissible indication of any teaching, suggestion or motivation to combine these references.

These three references span an extraordinarily large domain of interest. The Conklin Patent deals primarily with search queries. The Berger Application deals with data classification and makes no mention of identifying topics. The Bradford Patent deals with document classification and also makes no mention of identifying topics. Collectively, document classification, data classification and search queries cover a very broad range of topics. Defining such a large domain of interest serves notice that the Examiner is reaching to combine references that otherwise would not be permissibly combined, and is inconsistent with the notion that one skilled in the arts would reasonably combine references scattered across such a large domain of interest.

Furthermore, as is discussed further below, the Conklin Patent teaches a very different approach to identifying topics that conflicts with the approach presented in the present application. Therefore, it would not be reasonable for an individual skilled in the art to reasonably combine the Bradford and Berger Applications, which say nothing about the identification of topics, with the Conklin Patent to contend that their combination teaches identification of topics using the elements of the present application to render obvious the elements of independent claims 1, 16, 29, 39, 49 and 50.

For at least this reason there is no motivation to combine the Conklin Patent, the Bradford Application and the Berger Application. Therefore they can not be used as a basis of rejection. Reconsideration and allowance is respectfully requested of the claims.

***The Cited References Fail to Serve as a Basis for Prima Facie Rejection***

Assuming arguendo that the references could be combined, the combined references still fail to serve as a basis for a prima facie rejection. The Examiner admits that the Conklin Patent does not teach any of the elements of independent claims 1. Present Office Action at 4. The Examiner inappropriately uses the Conklin Patent in an attempt to impermissibly combine the disparate references by noting that Conklin teaches identifying topics of a data corpus of documents. The Examiner observes that Conklin teaches a method to identify topics. The Examiner provides support for this in the abstract of the Conklin Patent. Present Office Action at 4. The Examiner's reliance on the abstract is misplaced. In fact, the abstract of the Conklin Patent refers to matching topics that have already been assigned to documents to a query topic entered by a user.

This is an entirely different operation than identifying topics within a document or other data segment, which is the focus of the present application.

The Conklin Patent does subsequently address identifying topics within a document. In doing so, the Conklin Patent teaches away from the present invention by describing a process that is an entirely different way of determining topics. The Conklin Patent approach to identifying topics conflicts with the teachings of the present application.

To identify topics, the Conklin Patent teaches a complex approach that relies upon thematic tags. The thematic tags identify the thematic content of a document. Conklin Patent at Col. 14, Lines 63-65. Thematic tags are identified and associated with each document by a content processing system. Specifically, a linguistic engine extracts topics and content carrying words through use of thematic tags for each sentence in the documents within a data corpus. *Id.* at col. 14, lines 29-34. Furthermore, the Conklin Patent notes that "[e]ach word is discriminated in the text, identifying importance or meaning, the impact on different parts of the text, and the overall contribution to the content of the text. The thematic context of the text is determined in accordance with predetermined theme assessment criteria that is a function of the strategic importance of the discriminated words." *Id.* at col. 14, line 64 - col. 15, line 3. The Conklin Patent approach is entirely different and conflicts with the far simpler and more elegant approach within the present application that allows easy scaling. Unlike the Conklin Patent, the present application examines the frequency of word occurrence against an expected usage, without reference to word meanings or the need for predetermined theme assessment criteria. Given that the Conklin Patent teaches an entirely different

approach, which conflicts with the approach of the present application to identify topics, the Conklin Patent can not serve as a basis for rejecting the independent claims.

The Examiner also alleges that the Berger Patent "teaches actual value and expected value for each entry of words or pattern based on the frequency occurrences via threshold." Present Office Action at 4. In the support of this statement, the Examiner refers to paragraphs 0032-0033, 0042-0044. *Id.* The Berger Patent refers to the determination of an "interestingness measure," which is the measure of how much the occurrence of the pattern correlates with the occurrence of a single value of the predicted variable. Berger Patent at ¶ 0032. A significant difference between the approach in the Berger Patent and that in the present application is that the Berger Patent computes the expected usage value based on a previous data set, then computes the actual usage value of a pattern within a new data set and compares it to the existing expected usage value.

In the example given by the Berger Patent, for example, actual values for patterns during a malicious network event are compared with expected values for patterns during normal network activity to identify patterns that are non-frequent and yet highly correlated with intrusive behavior. *Id.* at ¶0033. This approach is at odds with the approach in the present application in that both the expected and actual values for patterns are determined using the same data set, thus no *a priori* knowledge or stored expected values are needed. This greatly simplifies the present approach. In fact using the Berger approach as explained with respect to the network intrusion example would not be applicable to identifying topics in a data corpus. Furthermore, the Berger Patent makes no mention of word combinations or using the approach specified in the Berger Patent to identify topics.



The Examiner attempts, in part, to rectify this shortcoming by combining the Berger Patent with the Bradford Patent. The Examiner contends that the Bradford Patent teaches contiguous words. Office Action at 5. The Bradford Patent, however, does not teach that contiguous words from word combinations that can be considered topics.

Numerous reasons have been provided above that show that the Conklin Patent, Berger Application, Bradford Application or their combination do not teach, disclose or suggest the elements of independent claims 1, 16, 29, 39, 49 and 50 of the present application.

Specifically, the Conklin Patent, the Berger Application and the Bradford Application do teach the element of designating a word combination as a topic if the segment level actual usage value of the word combination is greater than the segment-level expected usage value of the word combination.

Furthermore, with respect to claim 49, it includes the element of computing a segment-level expected usage value for each of the one or more word combinations, wherein the segment-level expected usage value is based on frequency counts of words that form the word combination within the data corpus or a portion thereof. This differs from independent claim 1 in that claim 49 includes the feature that the segment-level expected usage value is based on frequency counts of words that form the word combination. The Examiner has not discussed how the cited references disclose this feature of claim 49. In fact, none of the references suggest or teach this aspect of this feature.

For any one of the above reasons, the independent claims are patentable over the cited references. Reconsideration and allowance of the independent claims is respectfully requested.

Because each dependent claim incorporates all of the elements of the independent claim from which it depends, as well as additional features, the above arguments made with respect to the independent claims, apply *a fortiori* to the dependent claims. For at least this reason, the dependent claims are also patentable over the cited references and their combination. Reconsideration and allowance of these claims is respectfully requested.

Claims 29-48 and 50 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over the Conklin Patent in view of the Berger Application in further view of the Bradford Application. Office Action at 11. The Examiner's subsequent arguments, however, refer only to the Conklin Patent in combination with the Zimmerman Patent. Applicants assume the reference to the Berger Application and Bradford Application was a typographical error, and address the Examiner's arguments with respect to the Conklin and Zimmerman Patents.

The Zimmerman Patent does not address any of the shortcomings identified above with respect to the use of the Conklin Patent, Berger Application and Bradford Application to reject the claims in the present application.

As discussed in a previous office action response, the Zimmerman Patent discloses a document retrieval system that contains a database that relates document word-pair patterns to topics. Zimmerman Patent Abstract. The system is initialized through training using a small initial database that has been manually indexed such that each document in the training database is manually assigned to one or more pre-existing categories or topics. *Id.* at col. 7, lines 1-4. Initially, documents are assigned to pre-existing topics. *Id.* at col. 4, lines 64-65. New documents added to the system are

assigned topics by comparison of word pairing similarities in the new documents to the existing documents in the system. Topics are not word pairings that occur within a document, rather word pairings within a document are used to associate a document with a topic that is not the same as the word pairing. *Id.* at col. 14, lines 56-63. In fact within Zimmerman, many topics are identified as one word topics. *Id.* at col. 15, lines 20-25, *see also, Id.* at col. 4, line 52, *Id.* at Fig. 2. The Zimmerman Patent does not generate topics, nor is a word pairing in a document the same as a topic within the document. Unlike the Zimmerman Patent, the Present Application does generate topics and those topics are word combinations appearing in the documents that are indicated to have the particular topics.

There are numerous other fundamental differences between the Zimmerman Patent and the Present Application. For example, in the Present Application topics include word combinations that are two or more words that occur in the documents they are assigned to. In the Zimmerman Patent topics are most commonly one word topics, and the topics do not necessarily occur in the documents for which they are assigned. The word pairings identified in the Zimmerman Patent are not the same as topics. This contrasts with the approach in the Present Application in which the word combinations that meet the statistical tests become a topic.

Moreover, at the very least, the Zimmerman Patent does not teach, disclose or suggest the element of computing a segment-level expected usage value for a word combination element of the independent claims of the present application, nor does it teach, disclose or suggest the element of designating a word combination as a topic if the segment level usage value of the word combination is greater than the segment level usage value of the word combination element of the independent claims. Like the

Conklin Patent the Zimmerman Patent teaches away from the present application.

Combining two references that teach away from the present application, can not be used as a basis for rejecting the claims within the present application.

For any one of the above reasons, the independent claims are patentable over the cited references. Reconsideration and allowance of these independent claims is respectfully requested.

Because each dependent claim incorporates all of the elements of the independent claim from which it depends, as well as additional features, the above arguments made with respect to the independent claims, apply *a fortiori* to the dependent claims. For at least this reason, the dependent claims are also patentable over the cited references and their combination. Reconsideration and allowance of these claims is respectfully requested.

***Conclusion***

All of the stated grounds of objection and rejection have been properly traversed, accommodated, or rendered moot. Applicants therefore respectfully request that the Examiner reconsider all presently outstanding objections and rejections and that they be withdrawn. Applicants believe that a full and complete reply has been made to the outstanding Office Action and, as such, the present application is in condition for allowance. If the Examiner believes, for any reason, that personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided.

Prompt and favorable consideration of this Amendment and Reply is respectfully requested.

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

A handwritten signature in black ink, appearing to read "Michael D. Specht", written in a cursive style.

Michael D. Specht  
Attorney for Applicants  
Registration No. 54,463

Date: 3/1/06

1100 New York Avenue, N.W.  
Washington, D.C. 20005-3934  
(202) 371-2600